



SEQUENCE LISTING

RECEIVED

JUN 26 2001

TECH CENTER 1600/2900

<110> YAMAMOTO, Takuo
MARUTA, Kazuhiko
KUBOTA, Michio
FUKUDA, Shigeharu
MIYAKE, Toshio

<120> NON-REDUCING SACCHARIDE-FORMING ENZYME,
TREHALOSE-RELEASING ENZYME, AND PROCESS FOR PRODUCING
SACCHARIDES USING THE ENZYMES

<130> YAMAMOTO=16A

<140> 09/435,770

<141> 1999-11-08

<150> JP 258,394/1998

<151> 1998-09-11

<150> JP 352,252/1998

<151> 1998-12-11

<150> JP 16,931/1999

<151> 1999-01-26

<160> 39

<170> PatentIn Ver. 2.1

<210> 1

<211> 756

<212> PRT

<213> ARTHROBACTER sp.S34

<400> 1

Pro Ala Ser Thr Tyr Arg Leu Gln Ile Ser Ala Glu Phe Thr Leu Phe
1 5 10 15

Asp Ala Ala Arg Ile Val Pro Tyr Leu His Arg Leu Gly Ala Asp Trp
20 25 30

Leu Tyr Leu Ser Pro Leu Leu Glu Ser Glu Ser Gly Ser Ser His Gly
35 40 45

Tyr Asp Val Val Asp His Ser Arg Val Asp Ala Ala Arg Gly Gly Pro
50 55 60

Glu Gly Leu Ala Glu Leu Ser Arg Ala Ala His Glu Arg Gly Met Gly
65 70 75 80

Val Val Val Asp Ile Val Pro Asn His Val Gly Val Ala Thr Pro Lys
85 90 95

Ala Asn Arg Trp Trp Trp Asp Val Leu Ala Arg Gly Gln Arg Ser Glu
100 105 110

Tyr Ala Asp Tyr Phe Asp Ile Asp Trp Glu Phe Gly Gly Gly Arg Leu
115 120 125

Arg Leu Pro Val Leu Gly Asp Gly Pro Asp Glu Leu Asp Ala Leu Arg
130 135 140

Val Asp Gly Asp Glu Leu Val Tyr Tyr Glu His Arg Phe Pro Ile Ala
145 150 155 160

Glu Gly Thr Gly Gly Gly Thr Pro Arg Glu Val His Asp Arg Gln His
 165 170
 Tyr Glu Leu Met Ser Trp Arg Arg Ala Asp His Asp Leu Asn Tyr Arg
 180 185 190
 Arg Phe Phe Ala Val Asn Thr Leu Ala Ala Val Arg Val Glu Asp Pro
 195 200 205
 Arg Val Phe Asp Asp Thr His Arg Glu Ile Gly Arg Trp Ile Ala Glu
 210 215 220
 Gly Leu Val Asp Gly Leu Arg Val Asp His Pro Asp Gly Leu Arg Ala
 225 230 235 240
 Pro Gly Asp Tyr Leu Arg Arg Leu Ala Glu Leu Ala Gln Gly Arg Pro
 245 250 255
 Ile Trp Val Glu Lys Ile Ile Glu Gly Asp Glu Arg Met Pro Pro Gln
 260 265 270
 Trp Pro Ile Ala Gly Thr Thr Gly Tyr Asp Ala Leu Ala Gly Ile Asp
 275 280 285
 Arg Val Leu Val Asp Pro Ala Gly Glu His Pro Leu Thr Gln Ile Val
 290 295 300
 Asp Glu Ala Ala Gly Ser Pro Arg Arg Trp Ala Glu Leu Val Pro Glu
 305 310 315 320
 Arg Lys Arg Ala Val Ala Arg Gly Ile Leu Asn Ser Glu Ile Arg Arg
 325 330 335
 Val Ala Arg Glu Leu Gly Glu Val Ala Gly Asp Val Glu Asp Ala Leu
 340 345 350
 Val Glu Ile Ala Ala Ala Leu Ser Val Tyr Arg Ser Tyr Leu Pro Phe
 355 360 365
 Gly Arg Glu His Leu Asp Glu Ala Val Ala Ala Ala Gln Ala Ala Ala
 370 375 380
 Pro Gln Leu Glu Ala Asp Leu Ala Ala Val Gly Ala Ala Leu Ala Asp
 385 390 395 400
 Pro Gly Asn Pro Ala Ala Leu Arg Phe Gln Gln Thr Ser Gly Met Ile
 405 410 415
 Met Ala Lys Gly Val Glu Asp Asn Ala Phe Tyr Arg Tyr Pro Arg Leu
 420 425 430
 Thr Ser Leu Thr Glu Val Gly Gly Asp Pro Ser Leu Phe Ala Ile Asp
 435 440 445
 Ala Ala Ala Phe His Ala Ala Gln Arg Asp Arg Ala Ala Arg Leu Pro
 450 455 460
 Glu Ser Met Thr Thr Leu Thr Thr His Asp Thr Lys Arg Ser Glu Asp
 465 470 475 480
 Thr Arg Ala Arg Ile Thr Ala Leu Ala Glu Ala Pro Glu Arg Trp Arg
 485 490 495
 Arg Phe Leu Thr Glu Val Gly Gly Leu Ile Gly Thr Gly Asp Arg Val
 500 505 510

Leu Glu Asn Leu Ile Trp Gln Ala Ile Val Gly Ala Trp Pro Ala Ser
 515 520 525
 Arg Glu Arg Leu Glu Ala Tyr Ala Leu Lys Ala Ala Arg Glu Ala Gly
 530 535 540
 Glu Ser Thr Asp Trp Ile Asp Gly Asp Pro Ala Phe Glu Glu Arg Leu
 545 550 555 560
 Thr Arg Leu Val Thr Val Ala Val Glu Glu Pro Leu Val His Glu Leu
 565 570 575
 Leu Glu Arg Leu Val Asp Glu Leu Thr Ala Ala Gly Tyr Ser Asn Gly
 580 585 590
 Leu Ala Ala Lys Leu Leu Gln Leu Leu Ala Pro Gly Thr Pro Asp Val
 595 600 605
 Tyr Gln Gly Thr Glu Arg Trp Asp Arg Ser Leu Val Asp Pro Asp Asn
 610 615 620
 Arg Arg Pro Val Asp Phe Ala Ala Ala Ser Glu Leu Leu Asp Arg Leu
 625 630 635 640
 Asp Gly Gly Trp Arg Pro Pro Val Asp Glu Thr Gly Ala Val Lys Thr
 645 650 655
 Leu Val Val Ser Arg Ala Leu Arg Leu Arg Arg Asp Arg Pro Glu Leu
 660 665 670
 Phe Thr Ala Tyr His Pro Val Thr Ala Arg Gly Ala Gln Ala Glu His
 675 680 685
 Leu Ile Gly Phe Asp Arg Gly Gly Ala Ile Ala Leu Ala Thr Arg Leu
 690 695 700
 Pro Leu Gly Leu Ala Ala Ala Gly Gly Trp Gly Asp Thr Val Val Asp
 705 710 715 720
 Val Gly Glu Arg Ser Leu Arg Asp Glu Leu Thr Gly Arg Glu Ala Arg
 725 730 735
 Gly Ala Ala Arg Val Ala Glu Leu Phe Ala Asp Tyr Pro Val Ala Leu
 740 745 750
 Leu Val Glu Thr
 755

<210> 2
 <211> 6
 <212> PRT
 <213> ARTHROBACTER sp.S34

<400> 2
 Asp Ile Val Pro Asn His
 1 5

<210> 3
 <211> 6
 <212> PRT
 <213> ARTHROBACTER sp.S34

<400> 3
 Gly Thr Thr Gly Tyr Asp

1

5

<210> 4

<211> 20

<212> PRT

<213> ARTHROBACTER sp.S34

<400> 4

Pro Ala Ser Thr Tyr Arg Leu Gln Ile Ser Ala Glu Phe Thr Leu Phe
 1 5 10 15

Asp Ala Ala Arg
 20

<210> 5

<211> 20

<212> PRT

<213> ARTHROBACTER sp.S34

<400> 5

Ser Leu Val Asp Pro Asp Asn Arg Arg Pro Val Asp Phe Ala Ala Ala
 1 5 10 15

Ser Glu Leu Leu
 20

<210> 6

<211> 20

<212> PRT

<213> ARTHROBACTER sp.S34

<400> 6

Ala Asn Arg Trp Trp Trp Asp Val Leu Ala Arg Gly Gln Arg Ser Glu
 1 5 10 15

Tyr Ala Asp Tyr
 20

<210> 7

<211> 2268

<212> DNA

<213> ARTHROBACTER sp.S34

<400> 7

| | | | | | | |
|-------------|-------------|------------|-------------|------------|------------|------|
| cccgccagta | cctaccgcct | tcagatctcg | gcggagttca | ccctcttcga | cgcggcgcgc | 60 |
| atcgtgccct | acctgcaccg | cctcggcgcc | gactggctgt | acctctcgcc | gctgctcgag | 120 |
| tccgagtcgg | gctcctcgca | cggctacgac | gtggtcgacc | actccgcgct | cgacgccgcc | 180 |
| cgcggcgggc | cggaggggct | cgccgagctc | tcccgctgcg | cgcacgagcg | cgccatgggc | 240 |
| gtcgtcgtcg | acatcgtgcc | caaccacgtc | ggcgtcgcca | cgccgaaggc | gaaccgctgg | 300 |
| tggatgggacg | ttctggcccc | tggacagcgg | tcggagtacg | ccgactactt | cgacatcgac | 360 |
| tgggagttcg | gcggcggcag | gctgcgcctg | cccgtgctcg | gcgacggccc | cgacgagctc | 420 |
| gacgcgctga | gagtggatgg | cgacgagctc | gtctactacg | agcaccgctt | cccgatcgcc | 480 |
| gagggcaccg | gcggcggcac | cccgcgcgag | gtgcacgacc | ggcagcacta | cgagctgatg | 540 |
| tcgtggcggc | gggcccgaaca | cgacctcaac | taccgcccgt | tcttcgccgt | gaacacgctc | 600 |
| gccgccgtac | gcgtcgaaaga | ccgcgcgctg | ttcgacgaca | cccaccgcga | gatcggccgc | 660 |
| tggatcgccg | agggcctcgt | cgacggcctg | cgcgctcgacc | accccgacgg | gctgcgcgcc | 720 |
| cccggcgact | acctgcgcgc | tctcgccgag | ctcgcccaag | gcaggccgat | ctgggtcgag | 780 |
| aagatcatcg | agggcgacga | gcggatgccc | ccgcagtggc | ccatcgccgg | caccaccggc | 840 |
| tacgacgcgc | tggccgggat | cgaccgggtg | ctcgtcgacc | ccgcggggcg | gcacccgctc | 900 |
| acccagatcg | tcgacgaggg | ggcaggcagc | ccccggcgct | gggcccagct | ggttcccag | 960 |
| cgcaagcggg | ccgtcgcccc | cggcatcctg | aactccgaga | tccgcgcgct | cgcccgcgaa | 1020 |
| ctcggagagg | tcgcgcggcg | cgtcgaagac | gcgctcgctg | agatcgccgc | cgccctgtcc | 1080 |

```

gtctaccgca gctacctgcc gttcggggcg gagcacctcg acgaagccgt ggccgcccgc 1140
caggccgcag cccccagct cgaggccgac ctccgcccgc tcggcgcgag gctcgcgcac 1200
ccggggcaacc ccgcccgcgt ccgcttccag cagaccagcg gcatgatcat ggccaagggc 1260
gtcaggagaca acgcgtttcta ccgctacccc cggctcacct cgctgaccga ggtcggggga 1320
gacccgagcc tgttcgcgat cgacgcggcc gccttccacg cggcgcgagcg cgaccgcgcc 1380
gcccggctgc ccgagtcgat gacgacgctg accaccacg acaccaagcg cagcgaagac 1440
acccgggcgc ggatcacgcg gctcgcgcgag gccccgaac gctggcgggc cttcctgacc 1500
gaggtcggcg ggctcatcgg aacgggcgac cgggtgctgg agaacctgat ctggcaggcg 1560
atcgtcggcg cgtggccggc gagccgggag cggctcgagg cctacgcgct gaaggccgcg 1620
cgcgaaagccg gcgagtcgac cgactggatc gacggcgacc ccgcgttcga agagcggctg 1680
acccgcctgg tcacggtcgc cgtcgcaggag ccgctcgtgc acgagctgct cgagcggctc 1740
gtcgcagagc tgacggcggc cgggtactcc aacggcctcg cggcgaaagt gctgcagctg 1800
ctcgcctccg gaacccccga cgtgtaccag ggcacggaac gctgggaccg gtcgctggtg 1860
gacccggaca accgtcgcgc cgtggatttc gccgcggcat ccgagctgct cgaccgcctc 1920
gacggcggct ggcggccgcc cgtcgcagag accggcgcgg tcaagacgct cgtcgtctcc 1980
cgcgcgctgc ggctgcgcgc cgaccggccc gagctgttca ccgcgtacca cccggtcacg 2040
gcgcgcggcg cgcaggccga gcacctgac ggcttcgacc gcggcggcgc gatcgccctg 2100
gccacccgcc tgccgctcgg cctcgccgcc gcaggcggct ggggcgacac ggtcgtcgac 2160
gtcggcgagc ggagcctgcg cgacgagctg accggccgcg aggcgcgcgc 2220
gtggccgagt tgttcgcgca ctaccccgct gccctgctgg tggagaca 2268

```

<210> 8
 <211> 28
 <212> DNA
 <213> ARTHROBACTER sp.S34

<400> 8
 ttttttaata aaatcaggag gaaaaaat

28

<210> 9
 <211> 575
 <212> PRT
 <213> ARTHROBACTER sp.S34

<400> 9
 Met Asn Arg Arg Phe Pro Val Trp Ala Pro Gln Ala Ala Gln Val Thr
 1 5 10 15
 Leu Val Val Gly Gln Gly Arg Ala Glu Leu Pro Leu Thr Arg Asp Glu
 20 25 30
 Asn Gly Trp Trp Ala Leu Gln Gln Pro Trp Asp Gly Gly Pro Asp Leu
 35 40 45
 Val Asp Tyr Gly Tyr Leu Val Asp Gly Lys Gly Pro Phe Ala Asp Pro
 50 55 60
 Arg Ser Leu Arg Gln Pro Arg Gly Val His Glu Leu Gly Arg Glu Phe
 65 70 75 80
 Asp Pro Ala Arg Tyr Ala Trp Gly Asp Asp Gly Trp Arg Gly Arg Asp
 85 90 95
 Leu Thr Gly Ala Val Ile Tyr Glu Leu His Val Gly Thr Phe Thr Pro
 100 105 110
 Glu Gly Thr Leu Asp Ser Ala Ile Arg Arg Leu Asp His Leu Val Arg
 115 120 125
 Leu Gly Val Asp Ala Val Glu Leu Leu Pro Val Asn Ala Phe Asn Gly
 130 135 140
 Thr His Gly Trp Gly Tyr Asp Gly Val Leu Trp Tyr Ala Val His Glu
 145 150 155 160

Pro Tyr Gly Gly Pro Glu Ala Tyr Gln Arg Phe Val Asp Ala Cys His
 165 170 175
 Ala Arg Gly Leu Ala Val Val Gln Asp Val Val Tyr Asn His Leu Gly
 180 185 190
 Pro Ser Gly Asn His Leu Pro Asp Phe Gly Pro Tyr Leu Gly Ser Gly
 195 200 205
 Ala Ala Asn Thr Trp Gly Asp Ala Leu Asn Leu Asp Gly Pro Leu Ser
 210 215 220
 Asp Glu Val Arg Arg Tyr Ile Ile Asp Asn Ala Val Tyr Trp Leu Arg
 225 230 235 240
 Asp Met His Ala Asp Gly Leu Arg Leu Asp Ala Val His Ala Leu Arg
 245 250 255
 Asp Ala Arg Ala Leu His Leu Leu Glu Glu Leu Ala Ala Arg Val Asp
 260 265 270
 Glu Leu Ala Gly Glu Leu Gly Arg Pro Leu Thr Leu Ile Ala Glu Ser
 275 280 285
 Asp Leu Asn Asp Pro Lys Leu Ile Arg Ser Arg Ala Ala His Gly Tyr
 290 295 300
 Gly Leu Asp Ala Gln Trp Asp Asp Asp Val His His Ala Val His Ala
 305 310 315 320
 Asn Val Thr Gly Glu Thr Val Gly Tyr Tyr Ala Asp Phe Gly Gly Leu
 325 330 335
 Gly Ala Leu Val Lys Val Phe Gln Arg Gly Trp Phe His Asp Gly Thr
 340 345 350
 Trp Ser Ser Phe Arg Glu Arg His His Gly Arg Pro Leu Asp Pro Asp
 355 360 365
 Ile Pro Phe Arg Arg Leu Val Ala Phe Ala Gln Asp His Asp Gln Val
 370 375 380
 Gly Asn Arg Ala Val Gly Asp Arg Met Ser Ala Gln Val Gly Glu Gly
 385 390 395 400
 Ser Leu Ala Ala Ala Ala Ala Leu Val Leu Leu Gly Pro Phe Thr Pro
 405 410 415
 Met Leu Phe Met Gly Glu Glu Trp Gly Ala Arg Thr Pro Trp Gln Phe
 420 425 430
 Phe Thr Ser His Pro Glu Pro Glu Leu Gly Glu Ala Thr Ala Arg Gly
 435 440 445
 Arg Ile Ala Glu Phe Ala Arg Met Gly Trp Asp Pro Ala Val Val Pro
 450 455 460
 Asp Pro Gln Asp Pro Ala Thr Phe Ala Arg Ser His Leu Asp Trp Ser
 465 470 475 480
 Glu Pro Glu Arg Glu Pro His Ala Gly Leu Leu Ala Phe Tyr Thr Asp
 485 490 495
 Leu Ile Ala Leu Arg Arg Glu Leu Pro Val Asp Ala Pro Ala Arg Glu
 500 505 510

Val Asp Ala Asp Glu Ala Arg Gly Val Phe Ala Phe Ser Arg Gly Pro
515 520 525

Leu Arg Val Thr Val Ala Leu Arg Pro Gly Pro Val Gly Val Pro Glu
530 535 540

His Gly Gly Leu Val Leu Ala Tyr Gly Glu Val Arg Ala Gly Ala Ala
545 550 555 560

Gly Leu His Leu Asp Gly Pro Gly Ala Ala Ile Val Arg Leu Glu
565 570 575

<210> 10
<211> 6
<212> PRT
<213> ARTHROBACTER sp.S34

<400> 10
Trp Gly Tyr Asp Gly Val
1 5

<210> 11
<211> 6
<212> PRT
<213> ARTHROBACTER sp.S34

<400> 11
Asp Val Val Tyr Asn His
1 5

<210> 12
<211> 7
<212> PRT
<213> ARTHROBACTER sp.S34

<400> 12
Arg Leu Asp Ala Val His Ala
1 5

<210> 13
<211> 7
<212> PRT
<213> ARTHROBACTER sp.S34

<400> 13
Ile Ala Glu Ser Asp Leu Asn
1 5

<210> 14
<211> 20
<212> PRT
<213> ARTHROBACTER sp.S34

<400> 14
Met Asn Arg Arg Phe Pro Val Trp Ala Pro Gln Ala Ala Gln Val Thr
1 5 10 15

Leu Val Val Gly
20

<210> 15
 <211> 20
 <212> PRT
 <213> ARTHROBACTER sp.S34

<400> 15
 Ser Arg Ala Ala His Gly Tyr Gly Leu Asp Ala Gln Trp Asp Asp Asp
 1 5 10 15

Val His His Ala
 20

<210> 16
 <211> 20
 <212> PRT
 <213> ARTHROBACTER sp.S34

<400> 16
 Asp Glu Asn Gly Trp Trp Ala Leu Gln Gln Pro Trp Asp Gly Gly Pro
 1 5 10 15

Asp Leu Val Asp
 20

<210> 17
 <211> 1725
 <212> DNA
 <213> ARTHROBACTER sp.S34

<400> 17
 atgaaccgac gattccccggt ctgggcgccc caggccgcgc aggtgacgct cgctcgtgggc 60

| | | | | | | |
|-------------|------------|------------|------------|------------|-------------|------|
| caaggccgcg | cgaactccc | gctgacccgc | gacgagaacg | gatggtgggc | tcttcagcag | 120 |
| ccgtgggacg | gcggccccga | cctcgtcgac | tacggctacc | tcgtcgacgg | caaggggcccc | 180 |
| ttcgccgacc | cgcggtcgct | gcggcagccg | cgcggcgtgc | acgagctcgg | ccgcgaattc | 240 |
| gaccccgccc | gctacgcgtg | gggcgacgac | ggatggcgcg | gccgagacct | caccggagcc | 300 |
| gtgatctacg | aactgcacgt | cggcaccttc | accctgagg | gaacgctgga | cagcgccatc | 360 |
| cgtcgcctcg | accacctggt | gcgcctcggc | gtcgacgcgg | tcgagctgct | gcccgtcaac | 420 |
| gcgttcaacg | gcacccacgg | ctggggctac | gacggggtgc | tctggtacgc | ggtgcacgag | 480 |
| ccctacggcg | gcccggagge | gtaccagcgc | ttcgtcgacg | cctgccacgc | ccgcggcctc | 540 |
| gccgtcgtgc | aggacgtcgt | ctacaaccac | ctggggccga | gcggcaacca | cctgcccagc | 600 |
| ttcggccctt | acctcgggtc | gggcgcccgc | aacacctggg | gcgacgcgct | gaacctcgac | 660 |
| gggcgcgtct | ccgacgaggt | gcggcggtac | atcatcgaca | acgcggtgta | ctggctgcgc | 720 |
| gacatgcacg | ccgacgggct | gcggctcgac | gccgtgcacg | cgctgcgcga | cgcccgcgcg | 780 |
| ctgcacctgc | tcgaagagct | cgccgcccgc | gtcgacgagc | tggcgggcga | gctcggccgg | 840 |
| ccgctgacgc | tcatcgccga | gagcgacctg | aacgaccgga | agctgatccg | ctcccgcgcg | 900 |
| gcgcacggct | acggcctcga | cgcccagtgg | gacgacgacg | tgcaccacgc | ggtgcacgcc | 960 |
| aacgtgaccg | gcgagaccgt | cggctactac | gccgacttcg | gcgggctcgg | cgccctcgtc | 1020 |
| aaggtgttcc | agcgcggctg | gttccacgac | ggcacctggt | cgagcttccg | cgagcggcac | 1080 |
| cacggccggc | cgctcgaccc | cgacatcccc | ttccgcgggc | tcgtcgccct | cgcgcaggat | 1140 |
| cacgaccagg | tcggcaaccg | agcggtcggc | gaccgcatgt | cggcgcaggc | cggcgagggt | 1200 |
| tcgtcgcggc | cgcgggcggc | gctcgtgctg | ctcggcccgt | tcaccccgat | gctgttcattg | 1260 |
| ggcgaggagt | ggggcgcgcg | caccccgtgg | cagttcttca | cctcccaccc | cgagcccagag | 1320 |
| ctgggggagg | cgacggcgcg | cgggcgcatc | gccgagttcg | cccgcattgg | ctgggacccg | 1380 |
| gcagtcgtgc | ccgacccgca | ggacccggcc | accttcgccc | gctcgcacct | ggactggtcc | 1440 |
| gagcccagagc | gggaaccgca | cgcgggcctg | ctcgccttct | acaccgacct | gatcgcgctg | 1500 |
| cggcgcgagc | tgccggtcga | tgccgcccgg | cgcgaggtgg | atgccgacga | ggcgcgcggc | 1560 |
| gtcttcgcgt | tcagccgcgg | cccgtcgcg | gtcacggtcg | cgctgcgccc | cggaccggtc | 1620 |
| ggggtgcccc | agcacggggg | cctcgtgctc | gcctacggcg | aggtgcgcgc | cggcgccgcc | 1680 |
| ggactgcacc | tcgacggggc | gggagccgcg | atcgtgcgcc | tcgag | | 1725 |

<210> 18
 <211> 23

<212> DNA
 <213> ARTHROBACTER sp.S34

<400> 18
 gcsaaccgst ggtggtggga cgt

23

<210> 19
 <211> 3252
 <212> DNA
 <213> ARTHROBACTER sp.S34

<220>
 <221> 5'UTR
 <222> (1)..(742)

<220>
 <221> CDS
 <222> (743)..(3013)

<220>
 <221> 3'UTR
 <222> (742)..(3014)

<400> 19
 atgccgacga cgaacttgag cgcgttctcg ggcacccgcg agagcgggtcc ggcgacggcg 60
 gcgcccagtg ccacgacgag cagcatcgcg gcgagcgccg cgacgacggc gaccggcagg 120
 cgcccctgat tgctggcgaa ggtgagcagc atgaagacca cctcgaggcc ctcgagcaac 180
 acacctttga acgacacggt gaacgcgtac caatcggaga ccccgaaaccg gctctcgcg 240
 cgggcgctct cggccgcctc gacctgacgc cggaaggcag cctcctcgtc acggagagcc 300
 ctgcgccttg ccgcgcgcag caccgccttg cgcagccagc cgagcccgaa gacgagcagc 360
 aaccgcgccg cgacgaggcg cagcacggcc agcggcagca gcaggatcgc gggaccgacg 420
 agcgcgacgg ccgcggccag caccaccacg gcgacggcgg cacctgtcag cgccgaccgc 480
 cagctgcggg tggcgccgac cgcgacgacg atcgtggtcg cctccaccgc ctcgaccacg 540
 caggcgagga acacggcggc gaacaggggc acggcgggtca tcggcccagc agacgggttga 600
 ccatcacggc acgctagcgc cattgtcac aggaagggcc aagacgcccg caacgcggca 660
 cccgtggacg gcgcgtaccg gcgtgtgacc gatcgtgtca accggtggcg cccgcccga 720
 gcacctgcgt agattcggcc tc gtg ccc gcc agt acc tac cgc ctt cag atc 772
 Met Pro Ala Ser Thr Tyr Arg Leu Gln Ile
 1 5 10

tcg gcg gag ttc acc ctc ttc gac gcg gcg cgc atc gtg ccc tac ctg 820
 Ser Ala Glu Phe Thr Leu Phe Asp Ala Ala Arg Ile Val Pro Tyr Leu
 15 20 25

cac cgc ctc ggc gcc gac tgg ctg tac ctc tcg ccg ctg ctc gag tcc 868
 His Arg Leu Gly Ala Asp Trp Leu Tyr Leu Ser Pro Leu Leu Glu Ser
 30 35 40

gag tcg ggc tcc tcg cac ggc tac gac gtg gtc gac cac tcc cgc gtc 916
 Glu Ser Gly Ser Ser His Gly Tyr Asp Val Val Asp His Ser Arg Val
 45 50 55

gac gcc gcc cgc ggc ggg ccg gag ggg ctc gcc gag ctc tcc cgt gcg 964
 Asp Ala Ala Arg Gly Gly Pro Glu Gly Leu Ala Glu Leu Ser Arg Ala
 60 65 70

gcg cac gag cgc ggc atg ggc gtc gtc gtc gac atc gtg ccc aac cac 1012
 Ala His Glu Arg Gly Met Gly Val Val Val Asp Ile Val Pro Asn His
 75 80 85 90

gtc ggc gtc gcg acg ccg aag gcg aac cgc tgg tgg tgg gac gtt ctg 1060
 Val Gly Val Ala Thr Pro Lys Ala Asn Arg Trp Trp Trp Asp Val Leu
 95 100 105

| | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| gcc | cgt | gga | cag | cgg | tcg | gag | tac | gcc | gac | tac | ttc | gac | atc | gac | tgg | 1108 |
| Ala | Arg | Gly | Gln | Arg | Ser | Glu | Tyr | Ala | Asp | Tyr | Phe | Asp | Ile | Asp | Trp | |
| | | | 110 | | | | | 115 | | | | | 120 | | | |
| gag | ttc | ggc | ggc | ggc | agg | ctg | cgc | ctg | ccc | gtg | ctc | ggc | gac | ggc | ccc | 1156 |
| Glu | Phe | Gly | Gly | Gly | Arg | Leu | Arg | Leu | Pro | Val | Leu | Gly | Asp | Gly | Pro | |
| | | 125 | | | | | 130 | | | | | 135 | | | | |
| gac | gag | ctc | gac | gcg | ctg | aga | gtg | gat | ggc | gac | gag | ctc | gtc | tac | tac | 1204 |
| Asp | Glu | Leu | Asp | Ala | Leu | Arg | Val | Asp | Gly | Asp | Glu | Leu | Val | Tyr | Tyr | |
| | 140 | | | | | 145 | | | | | 150 | | | | | |
| gag | cac | cgc | ttc | ccg | atc | gcc | gag | ggc | acc | ggc | ggc | ggc | acc | ccg | cgc | 1252 |
| Glu | His | Arg | Phe | Pro | Ile | Ala | Glu | Gly | Thr | Gly | Gly | Gly | Thr | Pro | Arg | |
| 155 | | | | | 160 | | | | | 165 | | | | | 170 | |
| gag | gtg | cac | gac | cgg | cag | cac | tac | gag | ctg | atg | tcg | tgg | cgg | cgg | gcc | 1300 |
| Glu | Val | His | Asp | Arg | Gln | His | Tyr | Glu | Leu | Met | Ser | Trp | Arg | Arg | Ala | |
| | | | | 175 | | | | | 180 | | | | | 185 | | |
| gac | cac | gac | ctc | aac | tac | cgc | cgc | ttc | ttc | gcc | gtg | aac | acg | ctc | gcc | 1348 |
| Asp | His | Asp | Leu | Asn | Tyr | Arg | Arg | Phe | Phe | Ala | Val | Asn | Thr | Leu | Ala | |
| | | | 190 | | | | | 195 | | | | | 200 | | | |
| gcc | gta | cgc | gtc | gaa | gac | ccg | cgc | gtg | ttc | gac | gac | acc | cac | cgc | gag | 1396 |
| Ala | Val | Arg | Val | Glu | Asp | Pro | Arg | Val | Phe | Asp | Asp | Thr | His | Arg | Glu | |
| | | 205 | | | | | 210 | | | | | 215 | | | | |
| atc | ggc | cgc | tgg | atc | gcc | gag | ggc | ctc | gtc | gac | ggc | ctg | cgc | gtc | gac | 1444 |
| Ile | Gly | Arg | Trp | Ile | Ala | Glu | Gly | Leu | Val | Asp | Gly | Leu | Arg | Val | Asp | |
| | 220 | | | | | 225 | | | | | 230 | | | | | |
| cac | ccc | gac | ggg | ctg | cgc | gcc | ccc | ggc | gac | tac | ctg | cgc | cgt | ctc | gcc | 1492 |
| His | Pro | Asp | Gly | Leu | Arg | Ala | Pro | Gly | Asp | Tyr | Leu | Arg | Arg | Leu | Ala | |
| 235 | | | | | 240 | | | | | 245 | | | | | 250 | |
| gag | ctc | gcc | caa | ggc | agg | ccg | atc | tgg | gtc | gag | aag | atc | atc | gag | ggc | 1540 |
| Glu | Leu | Ala | Gln | Gly | Arg | Pro | Ile | Trp | Val | Glu | Lys | Ile | Ile | Glu | Gly | |
| | | | 255 | | | | | 260 | | | | | | 265 | | |
| gac | gag | cgg | atg | ccc | ccg | cag | tgg | ccc | atc | gcc | ggc | acc | acc | ggc | tac | 1588 |
| Asp | Glu | Arg | Met | Pro | Pro | Gln | Trp | Pro | Ile | Ala | Gly | Thr | Thr | Gly | Tyr | |
| | | | 270 | | | | | 275 | | | | | 280 | | | |
| gac | gcg | ctg | gcc | ggg | atc | gac | cgg | gtg | ctc | gtc | gac | ccc | gcg | ggc | gag | 1636 |
| Asp | Ala | Leu | Ala | Gly | Ile | Asp | Arg | Val | Leu | Val | Asp | Pro | Ala | Gly | Glu | |
| | | 285 | | | | | 290 | | | | | 295 | | | | |
| cat | ccg | ctc | acc | cag | atc | gtc | gac | gag | gcg | gca | ggc | agc | ccc | cgg | cgc | 1684 |
| His | Pro | Leu | Thr | Gln | Ile | Val | Asp | Glu | Ala | Ala | Gly | Ser | Pro | Arg | Arg | |
| | 300 | | | | | 305 | | | | | 310 | | | | | |
| tgg | gcc | gag | ctg | gtt | ccc | gag | cgc | aag | cgg | gcc | gtc | gcc | cgc | ggc | atc | 1732 |
| Trp | Ala | Glu | Leu | Val | Pro | Glu | Arg | Lys | Arg | Ala | Val | Ala | Arg | Gly | Ile | |
| 315 | | | | | 320 | | | | | 325 | | | | | 330 | |
| ctg | aac | tcc | gag | atc | cgc | cgc | gtc | gcc | cgc | gaa | ctc | gga | gag | gtc | gcc | 1780 |
| Leu | Asn | Ser | Glu | Ile | Arg | Arg | Val | Ala | Arg | Glu | Leu | Gly | Glu | Val | Ala | |
| | | | | 335 | | | | | 340 | | | | | 345 | | |
| ggc | gac | gtc | gaa | gac | gcg | ctc | gtc | gag | atc | gcc | gcc | gcc | ctg | tcc | gtc | 1828 |
| Gly | Asp | Val | Glu | Asp | Ala | Leu | Val | Glu | Ile | Ala | Ala | Ala | Leu | Ser | Val | |
| | | | 350 | | | | | 355 | | | | | 360 | | | |
| tac | cgc | agc | tac | ctg | ccg | ttc | ggg | cgc | gag | cac | ctc | gac | gaa | gcc | gtg | 1876 |
| Tyr | Arg | Ser | Tyr | Leu | Pro | Phe | Gly | Arg | Glu | His | Leu | Asp | Glu | Ala | Val | |

| 365 | | | | 370 | | | | 375 | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| gcc | gcc | gcg | cag | gcc | gca | gcc | ccc | cag | ctc | gag | gcc | gac | ctc | gcc | gcc | 1924 |
| Ala | Ala | Ala | Gln | Ala | Ala | Ala | Pro | Gln | Leu | Glu | Ala | Asp | Leu | Ala | Ala | |
| | 380 | | | | | 385 | | | | | 390 | | | | | |
| gtc | ggc | gca | gcg | ctc | gcc | gac | ccg | ggc | aac | ccc | gcc | gcg | ctc | cgc | ttc | 1972 |
| Val | Gly | Ala | Ala | Leu | Ala | Asp | Pro | Gly | Asn | Pro | Ala | Ala | Leu | Arg | Phe | |
| 395 | | | | | 400 | | | | | 405 | | | | | 410 | |
| cag | cag | acc | agc | ggc | atg | atc | atg | gcc | aag | ggc | gtc | gag | gac | aac | gcg | 2020 |
| Gln | Gln | Thr | Ser | Gly | Met | Ile | Met | Ala | Lys | Gly | Val | Glu | Asp | Asn | Ala | |
| | | | | 415 | | | | | 420 | | | | | 425 | | |
| ttc | tac | cgc | tac | ccc | cgg | ctc | acc | tcg | ctg | acc | gag | gtc | ggg | gga | gac | 2068 |
| Phe | Tyr | Arg | Tyr | Pro | Arg | Leu | Thr | Ser | Leu | Thr | Glu | Val | Gly | Gly | Asp | |
| | | | 430 | | | | | 435 | | | | | 440 | | | |
| ccg | agc | ctg | ttc | gcg | atc | gac | gcg | gcc | gcc | ttc | cac | gcg | gcg | cag | cgc | 2116 |
| Pro | Ser | Leu | Phe | Ala | Ile | Asp | Ala | Ala | Ala | Phe | His | Ala | Ala | Gln | Arg | |
| | | 445 | | | | | 450 | | | | | 455 | | | | |
| gac | cgc | gcc | gcc | cgg | ctg | ccc | gag | tcg | atg | acg | acg | ctg | acc | acc | cac | 2164 |
| Asp | Arg | Ala | Ala | Arg | Leu | Pro | Glu | Ser | Met | Thr | Thr | Leu | Thr | Thr | His | |
| | 460 | | | | | 465 | | | | | | 470 | | | | |
| gac | acc | aag | cgc | agc | gaa | gac | acc | cgg | gcg | cgg | atc | acc | gcg | ctc | gcc | 2212 |
| Asp | Thr | Lys | Arg | Ser | Glu | Asp | Thr | Arg | Ala | Arg | Ile | Thr | Ala | Leu | Ala | |
| 475 | | | | | 480 | | | | | 485 | | | | | 490 | |
| gag | gcc | ccc | gaa | cgc | tgg | cgg | cgc | ttc | ctg | acc | gag | gtc | ggc | ggg | ctc | 2260 |
| Glu | Ala | Pro | Glu | Arg | Trp | Arg | Arg | Phe | Leu | Thr | Glu | Val | Gly | Gly | Leu | |
| | | | | 495 | | | | | 500 | | | | | 505 | | |
| atc | gga | acg | ggc | gac | cgg | gtg | ctg | gag | aac | ctg | atc | tgg | cag | gcg | atc | 2308 |
| Ile | Gly | Thr | Gly | Asp | Arg | Val | Leu | Glu | Asn | Leu | Ile | Trp | Gln | Ala | Ile | |
| | | | 510 | | | | | 515 | | | | | 520 | | | |
| gtc | ggc | gcg | tgg | ccg | gcg | agc | cgg | gag | cgg | ctc | gag | gcc | tac | gcg | ctg | 2356 |
| Val | Gly | Ala | Trp | Pro | Ala | Ser | Arg | Glu | Arg | Leu | Glu | Ala | Tyr | Ala | Leu | |
| | | 525 | | | | | 530 | | | | | 535 | | | | |
| aag | gcc | gcg | cgc | gaa | gcc | ggc | gag | tcg | acc | gac | tgg | atc | gac | ggc | gac | 2404 |
| Lys | Ala | Ala | Arg | Glu | Ala | Gly | Glu | Ser | Thr | Asp | Trp | Ile | Asp | Gly | Asp | |
| | 540 | | | | | 545 | | | | | 550 | | | | | |
| ccc | gcg | ttc | gaa | gag | cgg | ctg | acc | cgc | ctg | gtc | acg | gtc | gcc | gtc | gag | 2452 |
| Pro | Ala | Phe | Glu | Glu | Arg | Leu | Thr | Arg | Leu | Val | Thr | Val | Ala | Val | Glu | |
| 555 | | | | | 560 | | | | | 565 | | | | | 570 | |
| gag | ccg | ctc | gtg | cac | gag | ctg | ctc | gag | cgg | ctc | gtc | gac | gag | ctg | acg | 2500 |
| Glu | Pro | Leu | Val | His | Glu | Leu | Leu | Glu | Arg | Leu | Val | Asp | Glu | Leu | Thr | |
| | | | | 575 | | | | | 580 | | | | | 585 | | |
| gcg | gcc | ggg | tac | tcc | aac | ggc | ctc | gcg | gcg | aag | ctg | ctg | cag | ctg | ctc | 2548 |
| Ala | Ala | Gly | Tyr | Ser | Asn | Gly | Leu | Ala | Ala | Lys | Leu | Leu | Gln | Leu | Leu | |
| | | | 590 | | | | | 595 | | | | | 600 | | | |
| gcc | ccc | gga | acc | ccc | gac | gtg | tac | cag | ggc | acg | gaa | cgc | tgg | gac | cgg | 2596 |
| Ala | Pro | Gly | Thr | Pro | Asp | Val | Tyr | Gln | Gly | Thr | Glu | Arg | Trp | Asp | Arg | |
| | | 605 | | | | | 610 | | | | | 615 | | | | |
| tcg | ctg | gtg | gac | ccg | gac | aac | cgt | cgc | ccc | gtg | gat | ttc | gcc | gcg | gca | 2644 |
| Ser | Leu | Val | Asp | Pro | Asp | Asn | Arg | Arg | Pro | Val | Asp | Phe | Ala | Ala | Ala | |
| | 620 | | | | | 625 | | | | | 630 | | | | | |

| | |
|---|------|
| tcc gag ctg ctc gac cgc ctc gac ggc ggc tgg cgg ccg ccc gtc gac | 2692 |
| Ser Glu Leu Leu Asp Arg Leu Asp Gly Gly Trp Arg Pro Pro Val Asp | |
| 635 640 645 650 | |
| gag acc ggc gcg gtc aag acg ctc gtc gtc tcc cgc gcg ctg cgg ctg | 2740 |
| Glu Thr Gly Ala Val Lys Thr Leu Val Val Ser Arg Ala Leu Arg Leu | |
| 655 660 665 | |
| cgc cgc gac cgg ccc gag ctg ttc acc gcg tac cac ccg gtc acg gcg | 2788 |
| Arg Arg Asp Arg Pro Glu Leu Phe Thr Ala Tyr His Pro Val Thr Ala | |
| 670 675 680 | |
| cgc ggc gcg cag gcc gag cac ctg atc ggc ttc gac cgc ggc ggc gcg | 2836 |
| Arg Gly Ala Gln Ala Glu His Leu Ile Gly Phe Asp Arg Gly Gly Ala | |
| 685 690 695 | |
| atc gcc ctg gcc acc cgc ctg ccg ctc ggc ctc gcc gcc gca ggc ggc | 2884 |
| Ile Ala Leu Ala Thr Arg Leu Pro Leu Gly Leu Ala Ala Ala Gly Gly | |
| 700 705 710 | |
| tgg ggc gac acg gtc gtc gac gtc ggc gag cgg agc ctg cgc gac gag | 2932 |
| Trp Gly Asp Thr Val Val Asp Val Gly Glu Arg Ser Leu Arg Asp Glu | |
| 715 720 725 730 | |
| ctg acc ggc cgc gag gcc cgc gga gcg gcg cgc gtg gcc gag ttg ttc | 2980 |
| Leu Thr Gly Arg Glu Ala Arg Gly Ala Ala Arg Val Ala Glu Leu Phe | |
| 735 740 745 | |
| gcc gac tac ccc gtc gcc ctg ctg gtg gag aca tgaaccgacg attcccggtc | 3033 |
| Ala Asp Tyr Pro Val Ala Leu Leu Val Glu Thr | |
| 750 755 | |
| tgggcgcccc aggcgcgcga ggtgacgctc gtcgtgggccc aaggcgcgcg cgaactcccg | 3093 |
| ctgacccgcg acgagaacgg atgggtgggct cttcagcagc cgtggggacgg cggccccgac | 3153 |
| ctcgtcgact acggctacct cgtcgacggc aagggccctt tcgccgaccc gcggtcgctg | 3213 |
| cggcagccgc gcggcgtgca cgagctcggc cgcgaattc | 3252 |

<210> 20
 <211> 26
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:SYNTHETIC

<400> 20
 atgcccgcga gtacctaccg ccttca 26

<210> 21
 <211> 25
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:SYNTHETIC

<400> 21
 tcatgtctcc accagcaggg cgacg 25

<210> 22
 <211> 50
 <212> DNA

<213> Artificial Sequence
 <220>
 <223> Description of Artificial Sequence:SYNTHETIC
 <400> 22
 aattcttttt taataaaatc aggaggaatc tagatgttta ctagtctgca 50
 <210> 23
 <211> 42
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Description of Artificial Sequence:SYNTHETIC
 <400> 23
 gactagtaaa catctagatt cctcctgatt ttattaaaaa ag 42
 <210> 24
 <211> 33
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Description of Artificial Sequence:SYNTHETIC
 <400> 24
 aaatctagat gcccgccagt acctaccgcc ttc 33
 <210> 25
 <211> 33
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Description of Artificial Sequence:SYNTHETIC
 <400> 25
 aaaactagtt tatcatgtct ccaccagcag ggc 33
 <210> 26
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Description of Artificial Sequence:SYNTHETIC
 <400> 26
 atcgtgatg tcggcgatat ag 22
 <210> 27
 <211> 29
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Description of Artificial Sequence:SYNTHETIC
 <400> 27
 gtactggcgg gcatatTTTT tcctcctga 29

<210> 28
 <211> 31
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence:SYNTHETIC

 <400> 28
 aatcaggagg aaaaaatatg cccgccagta c 31

 <210> 29
 <211> 22
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence:SYNTHETIC

 <400> 29
 tcgacgatct gggtagcg at 22

 <210> 30
 <211> 22
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence:SYNTHETIC

 <400> 30
 tcgacgagca cccggtcgat cc 22

 <210> 31
 <211> 26
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence:SYNTHETIC

 <220>
 <223> n at position 18 is a, c, g, or t.

 <221> misc_fea
 <222> (18)..(18)

 <400> 31
 cartggayg aygaygtnc a ycaygc 26

 <210> 32
 <211> 2218
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence:SYNTHETIC

 <220>
 <221> CDS

<222> (477)..(2201)

<220>

<221> 3'UTR

<222> (2202)..(2218)

<400> 32

```
ctgcagctgc tcgcccccg aacccccgac gtgtaccagg gcacggaacg ctgggaccgg      60
tcgctggtgg acccggaaca ccgtcgcccc gtggatttcg ccgcggcatc cgagctgctc     120
gaccgcctcg acggcggctg gcggccgccc gtcgacgaga ccggcgcggt caagacgctc     180
gtcgtctccc gcgcgtgctg gctgcgccgc gaccggcccc agctgttcac cgcgtaccac     240
ccggtcacgg cgcgcgggcg gcaggccgag cacctgatcg gcttcgaccg cggcggcgcg      300
atcgccctgg ccacccgcct gccgctcggc ctcgccgccg caggcggctg gggcgacacg     360
gtcgtcgacg tcggcgagcg gagcctgcgc gacgagctga ccggccgcga ggcccccgga     420
gcggcgcgcg tggccgagtt gttcgccgac taccctgctg ccctgctggt ggagac atg      479
                                         Met
                                         1
```

```
aac cga cga ttc ccg gtc tgg gcg ccc cag gcc gcg cag gtg acg ctc      527
Asn Arg Arg Phe Pro Val Trp Ala Pro Gln Ala Ala Gln Val Thr Leu
                    5                                10                                15
```

```
gtc gtg ggc caa ggc cgc gcc gaa ctc ccg ctg acc cgc gac gag aac      575
Val Val Gly Gln Gly Arg Ala Glu Leu Pro Leu Thr Arg Asp Glu Asn
                    20                                25                                30
```

```
gga tgg tgg gct ctt cag cag ccg tgg gac ggc ggc ccc gac ctc gtc      623
Gly Trp Trp Ala Leu Gln Gln Pro Trp Asp Gly Gly Pro Asp Leu Val
                    35                                40                                45
```

```
gac tac ggc tac ctc gtc gac ggc aag ggc ccc ttc gcc gac ccg cgg      671
Asp Tyr Gly Tyr Leu Val Asp Gly Lys Gly Pro Phe Ala Asp Pro Arg
                    50                                55                                60                                65
```

```
tcg ctg ccg cag ccg cgc ggc gtg cac gag ctc ggc cgc gaa ttc gac      719
Ser Leu Arg Gln Pro Arg Gly Val His Glu Leu Gly Arg Glu Phe Asp
                    70                                75                                80
```

```
ccc gcc cgc tac gcg tgg ggc gac gac gga tgg cgc ggc cga gac ctc      767
Pro Ala Arg Tyr Ala Trp Gly Asp Asp Gly Trp Arg Gly Arg Asp Leu
                    85                                90                                95
```

```
acc gga gcc gtg atc tac gaa ctg cac gtc ggc acc ttc acc cct gag      815
Thr Gly Ala Val Ile Tyr Glu Leu His Val Gly Thr Phe Thr Pro Glu
                    100                                105                                110
```

```
gga acg ctg gac agc gcc atc cgt cgc ctc gac cac ctg gtg cgc ctc      863
Gly Thr Leu Asp Ser Ala Ile Arg Arg Leu Asp His Leu Val Arg Leu
                    115                                120                                125
```

```
ggc gtc gac gcg gtc gag ctg ctg ccc gtc aac gcg ttc aac ggc acc      911
Gly Val Asp Ala Val Glu Leu Leu Pro Val Asn Ala Phe Asn Gly Thr
                    130                                135                                140                                145
```

```
cac ggc tgg ggc tac gac ggg gtg ctc tgg tac gcg gtg cac gag ccc      959
His Gly Trp Gly Tyr Asp Gly Val Leu Trp Tyr Ala Val His Glu Pro
                    150                                155                                160
```

```
tac ggc ggc ccg gag gcg tac cag cgc ttc gtc gac gcc tgc cac gcc      1007
```

| | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|--|
| Tyr | Gly | Gly | Pro | Glu | Ala | Tyr | Gln | Arg | Phe | Val | Asp | Ala | Cys | His | Ala | | |
| | | | 165 | | | | | 170 | | | | | 175 | | | | |
| cgc | ggc | ctc | gcc | gtc | gtg | cag | gac | gtc | gtc | tac | aac | cac | ctg | ggc | ccg | 1055 | |
| Arg | Gly | Leu | Ala | Val | Val | Gln | Asp | Val | Val | Tyr | Asn | His | Leu | Gly | Pro | | |
| | | 180 | | | | | 185 | | | | | 190 | | | | | |
| agc | ggc | aac | cac | ctg | ccc | gac | ttc | ggc | ccc | tac | ctc | ggg | tcg | ggc | gcc | 1103 | |
| Ser | Gly | Asn | His | Leu | Pro | Asp | Phe | Gly | Pro | Tyr | Leu | Gly | Ser | Gly | Ala | | |
| | 195 | | | | | 200 | | | | | 205 | | | | | | |
| gcc | aac | acc | tgg | ggc | gac | gcg | ctg | aac | ctc | gac | ggg | ccg | ctc | tcc | gac | 1151 | |
| Ala | Asn | Thr | Trp | Gly | Asp | Ala | Leu | Asn | Leu | Asp | Gly | Pro | Leu | Ser | Asp | | |
| 210 | | | | | 215 | | | | | 220 | | | | | 225 | | |
| gag | gtg | cgg | cgg | tac | atc | atc | gac | aac | gcg | gtg | tac | tgg | ctg | cgc | gac | 1199 | |
| Glu | Val | Arg | Arg | Tyr | Ile | Ile | Asp | Asn | Ala | Val | Tyr | Trp | Leu | Arg | Asp | | |
| | | | | 230 | | | | | 235 | | | | | 240 | | | |
| atg | cac | gcc | gac | ggg | ctg | cgg | ctc | gac | gcc | gtg | cac | gcg | ctg | cgc | gac | 1247 | |
| Met | His | Ala | Asp | Gly | Leu | Arg | Leu | Asp | Ala | Val | His | Ala | Leu | Arg | Asp | | |
| | | | 245 | | | | | 250 | | | | | 255 | | | | |
| gcc | cgc | gcg | ctg | cac | ctg | ctc | gaa | gag | ctc | gcc | gcc | cgc | gtc | gac | gag | 1295 | |
| Ala | Arg | Ala | Leu | His | Leu | Leu | Glu | Glu | Leu | Ala | Ala | Arg | Val | Asp | Glu | | |
| | | 260 | | | | | 265 | | | | | 270 | | | | | |
| ctg | gcg | ggc | gag | ctc | ggc | cgg | ccg | ctg | acg | ctc | atc | gcc | gag | agc | gac | 1343 | |
| Leu | Ala | Gly | Glu | Leu | Gly | Arg | Pro | Leu | Thr | Leu | Ile | Ala | Glu | Ser | Asp | | |
| | 275 | | | | | 280 | | | | | 285 | | | | | | |
| ctg | aac | gac | ccg | aag | ctg | atc | cgc | tcc | cgc | gcg | gcg | cac | ggc | tac | ggc | 1391 | |
| Leu | Asn | Asp | Pro | Lys | Leu | Ile | Arg | Ser | Arg | Ala | Ala | His | Gly | Tyr | Gly | | |
| 290 | | | | | 295 | | | | 300 | | | | | 305 | | | |
| ctc | gac | gcc | cag | tgg | gac | gac | gac | gtg | cac | cac | gcg | gtg | cac | gcc | aac | 1439 | |
| Leu | Asp | Ala | Gln | Trp | Asp | Asp | Asp | Val | His | His | Ala | Val | His | Ala | Asn | | |
| | | | | 310 | | | | | 315 | | | | | 320 | | | |
| gtg | acc | ggc | gag | acc | gtc | ggc | tac | tac | gcc | gac | ttc | ggc | ggg | ctc | ggc | 1487 | |
| Val | Thr | Gly | Glu | Thr | Val | Gly | Tyr | Tyr | Ala | Asp | Phe | Gly | Gly | Leu | Gly | | |
| | | | 325 | | | | | 330 | | | | | 335 | | | | |
| gcc | ctc | gtc | aag | gtg | ttc | cag | cgc | ggc | tgg | ttc | cac | gac | ggc | acc | tgg | 1535 | |
| Ala | Leu | Val | Lys | Val | Phe | Gln | Arg | Gly | Trp | Phe | His | Asp | Gly | Thr | Trp | | |
| | | 340 | | | | | 345 | | | | | 350 | | | | | |
| tcg | agc | ttc | cgc | gag | cgg | cac | cac | ggc | cgg | ccg | ctc | gac | ccc | gac | atc | 1583 | |
| Ser | Ser | Phe | Arg | Glu | Arg | His | His | Gly | Arg | Pro | Leu | Asp | Pro | Asp | Ile | | |
| | 355 | | | | | 360 | | | | | 365 | | | | | | |
| ccg | ttc | cgc | cgg | ctc | gtc | gcc | ttc | gcg | cag | gat | cac | gac | cag | gtc | ggc | 1631 | |
| Pro | Phe | Arg | Arg | Leu | Val | Ala | Phe | Ala | Gln | Asp | His | Asp | Gln | Val | Gly | | |
| 370 | | | | | 375 | | | | | 380 | | | | | 385 | | |
| aac | cga | gcg | gtc | ggc | gac | cgc | atg | tcg | gcg | cag | gtc | ggc | gag | ggt | tcg | 1679 | |
| Asn | Arg | Ala | Val | Gly | Asp | Arg | Met | Ser | Ala | Gln | Val | Gly | Glu | Gly | Ser | | |
| | | | | 390 | | | | | 395 | | | | 400 | | | | |
| ctc | gcc | gcc | gcg | gcg | gcg | ctc | gtg | ctg | ctc | ggc | ccg | ttc | acc | ccg | atg | 1727 | |
| Leu | Ala | Ala | Ala | Ala | Ala | Ala | Leu | Val | Leu | Gly | Pro | Phe | Thr | Pro | Met | | |
| | | | 405 | | | | | 410 | | | | | 415 | | | | |
| ctg | ttc | atg | ggc | gag | gag | tgg | ggc | gcg | cgc | acc | ccg | tgg | cag | ttc | ttc | 1775 | |
| Leu | Phe | Met | Gly | Glu | Glu | Trp | Gly | Ala | Arg | Thr | Pro | Trp | Gln | Phe | Phe | | |
| | | 420 | | | | | 425 | | | | | 430 | | | | | |

| | |
|---|------|
| acc tcc cac ccc gag ccc gag ctg ggg gag gcg acg gcg cgc ggg cgc | 1823 |
| Thr Ser His Pro Glu Pro Glu Leu Gly Glu Ala Thr Ala Arg Gly Arg | |
| 435 440 445 | |
| | |
| atc gcc gag ttc gcc cgc atg ggc tgg gac ccg gca gtc gtg ccc gac | 1871 |
| Ile Ala Glu Phe Ala Arg Met Gly Trp Asp Pro Ala Val Val Pro Asp | |
| 450 455 460 465 | |
| | |
| ccg cag gac ccg gcc acc ttc gcc cgc tcg cac ctg gac tgg tcc gag | 1919 |
| Pro Gln Asp Pro Ala Thr Phe Ala Arg Ser His Leu Asp Trp Ser Glu | |
| 470 475 480 | |
| | |
| ccc gag cgg gaa ccg cac gcg ggc ctg ctc gcc ttc tac acc gac ctg | 1967 |
| Pro Glu Arg Glu Pro His Ala Gly Leu Leu Ala Phe Tyr Thr Asp Leu | |
| 485 490 495 | |
| | |
| atc gcg ctg cgg cgc gag ctg ccg gtc gat gcg ccg gcg cgc gag gtg | 2015 |
| Ile Ala Leu Arg Arg Glu Leu Pro Val Asp Ala Pro Ala Arg Glu Val | |
| 500 505 510 | |
| | |
| gat gcc gac gag gcg cgc ggc gtc ttc gcg ttc agc cgc ggc ccg ctg | 2063 |
| Asp Ala Asp Glu Ala Arg Gly Val Phe Ala Phe Ser Arg Gly Pro Leu | |
| 515 520 525 | |
| | |
| cgg gtc acg gtc gcg ctg cgc ccc gga ccg gtc ggg gtg ccc gag cac | 2111 |
| Arg Val Thr Val Ala Leu Arg Pro Gly Pro Val Gly Val Pro Glu His | |
| 530 535 540 545 | |
| | |
| ggg ggc ctc gtg ctc gcc tac ggc gag gtg cgc gcc ggc gcc gcc gga | 2159 |
| Gly Gly Leu Val Leu Ala Tyr Gly Glu Val Arg Ala Gly Ala Ala Gly | |
| 550 555 560 | |
| | |
| ctg cac ctc gac ggg ccg gga gcc gcg atc gtg cgc ctc gag | 2201 |
| Leu His Leu Asp Gly Pro Gly Ala Ala Ile Val Arg Leu Glu | |
| 565 570 575 | |
| | |
| tgacgcggct gggtacc | 2218 |

<210> 33
 <211> 25
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:SYNTHETIC

<400> 33
 atgaaccgac gattcccggt ctggg 25

<210> 34
 <211> 25
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:SYNTHETIC

<400> 34
 tcactcgagg cgcacgatcg cggct 25

<210> 35
 <211> 36
 <212> DNA
 <213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:SYNTHETIC

<400> 35
aaatctagat gaaccgacga ttcccgggtct gggcgc 36

<210> 36
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:SYNTHETIC

<400> 36
aaaactagtt tatcactcga ggcgacgat cgcggc 36

<210> 37
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:SYNTHETIC

<400> 37
atcgtcgggt catatTTTTT cctcctga 28

<210> 38
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:SYNTHETIC

<400> 38
aatcaggagg aaaaaatatg aaccgacg 28

<210> 39
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:SYNTHETIC

<400> 39
agtggttgt agacgacgtc ct 22

C1
Cont